



TITLE:

Neural coding of syntactic structure in learned vocalizations in the songbird.

AUTHOR(S):

Fujimoto, Hisataka; Hasegawa, Taku; Watanabe, Dai

CITATION:

Fujimoto, Hisataka ...[et al]. Neural coding of syntactic structure in learned vocalizations in the songbird.. The Journal of neuroscience : the official journal of the Society for Neuroscience 2011, 31(27): 10023-10033

ISSUE DATE:

2011-07-06

URL:

<http://hdl.handle.net/2433/142953>

RIGHT:

©2011 the authors 0270-6474/11/3110023-11\$15.00/0; 許諾条件により本文は2012-01-06に公開.

Neural Coding of Syntactic Structure in Learned Vocalizations in the Songbird

Hisataka Fujimoto,¹ Taku Hasegawa,² and Dai Watanabe^{1,2}

¹Department of Biological Sciences, Faculty of Medicine and ²Department of Molecular and Systems Biology, Graduate School of Biostudies, Kyoto University, Kyoto, Kyoto 606-8501, Japan

Although vocal signals including human languages are composed of a finite number of acoustic elements, complex and diverse vocal patterns can be created from combinations of these elements, linked together by syntactic rules. To enable such syntactic vocal behaviors, neural systems must extract the sequence patterns from auditory information and establish syntactic rules to generate motor commands for vocal organs. However, the neural basis of syntactic processing of learned vocal signals remains largely unknown. Here we report that the basal ganglia projecting premotor neurons (HVC_X neurons) in Bengalese finches represent syntactic rules that generate variable song sequences. When vocalizing an alternative transition segment between song elements called syllables, sparse burst spikes of HVC_X neurons code the identity of a specific syllable type or a specific transition direction among the alternative trajectories. When vocalizing a variable repetition sequence of the same syllable, HVC_X neurons not only signal the initiation and termination of the repetition sequence but also indicate the progress and state-of-completeness of the repetition. These different types of syntactic information are frequently integrated within the activity of single HVC_X neurons, suggesting that syntactic attributes of the individual neurons are not programmed as a basic cellular subtype in advance but acquired in the course of vocal learning and maturation. Furthermore, some auditory–vocal mirroring type HVC_X neurons display transition selectivity in the auditory phase, much as they do in the vocal phase, suggesting that these songbirds may extract syntactic rules from auditory experience and apply them to form their own vocal behaviors.

Introduction

As a fundamental interface for communication, vocal signals convey information in their acoustic and sequence patterns. While vocal signals, including human languages, are composed of a finite set of acoustic elements, syntactic rules expand the diversity of sequence patterns (Chomsky, 1965; Bolhuis et al., 2010). To enable such syntactic vocal communication, neural systems must extract the sequence patterns from auditory information and establish syntactic rules to generate motor commands for vocal sequences (Chomsky, 1969; Santelmann and Jusczyk, 1998; Saffran, 2002; Kuhl, 2004). However, the neural processing of syntax for learned vocal signals remains largely unknown.

Bengalese finches assemble a fixed number of vocal elements called syllables into variable sequences (Woolley and Rubel, 1997) based on a finite-state type of syntax (Okanoya, 2004).

These songbirds, like humans, establish and maintain variable songs under the guidance of auditory information (Konishi, 1965; Brainard and Doupe, 2000). Therefore, the brain circuits involved in both perception and expression of vocal signals should contain the neurons that transmit syntactic information. In the songbird brain, telencephalic nucleus HVC (used as a proper name) functions in both sensory and motor processing of songs (Nottebohm et al., 1976; Vu et al., 1994; Yu and Margoliash, 1996; Gentner et al., 2000; Hahnloser et al., 2002; Long and Fee, 2008). Hence, analysis of neural activity in Bengalese finch HVC is expected to provide fundamental insights into the neural basis of syntactic organization.

HVC directly and indirectly receives auditory input (Kelley and Nottebohm, 1979; Katz and Gurney, 1981; McCasland and Konishi, 1981; Margoliash, 1983). Information processed in HVC is transmitted by two segregated types of projection neurons (Dutar et al., 1998; Mooney, 2000): HVC_{RA} neurons that project to the robust nucleus of the arcopallium (RA), which relay the information to motor neurons for vocal organs (Nottebohm et al., 1976); and HVC_X neurons, which innervate to basal ganglia area X (Nottebohm et al., 1982) and are involved in audition-dependent vocal plasticity (Scharff and Nottebohm, 1991; Brainard and Doupe, 2000). Both types of projection neurons generate sparse burst spikes during singing (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Prather et al., 2008). In the awake nonsinging state, however, HVC_{RA} neurons are completely inactive (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Prather et al., 2008). In contrast, a subpopulation of HVC_X neurons exhibits both motor-related activity and auditory re-

Received March 30, 2011; revised May 13, 2011; accepted May 26, 2011.

Author contributions: H.F. and D.W. designed research; H.F. performed research; H.F. and T.H. contributed unpublished reagents/analytic tools; H.F. analyzed data; H.F. and D.W. wrote the paper.

This work was supported by research grants from the Takeda Science Foundation, the Uehara Memorial Foundation, and the Ministry of Education, Culture, Sports, Science, and Technology of Japan including the Strategic Research Program for Brain Sciences. We thank S. Nakanishi and M. Konishi for discussions, advice, and criticisms that greatly benefited this project. We thank K. Tashiro for experimental assistance; A. Takashima for developing custom-made microdrives; Y. Funabiki and T. Hayakawa for developing tutoring methods; and R. Yamada, D. Koketsu, and A. Nambu for valuable discussion. We thank J. Hejna for careful reading and criticisms.

Correspondence should be addressed to Dr. Dai Watanabe, Department of Biological Sciences, Faculty of Medicine, Department of Molecular and Systems Biology, Graduate School of Biostudies, Kyoto University, Yoshida, Sakyo-ku, Kyoto 606-8501, Japan. E-mail: dai@phy.med.kyoto-u.ac.jp.

DOI:10.1523/JNEUROSCI.1606-11.2011

Copyright © 2011 the authors 0270-6474/11/3110023-11\$15.00/0

sponse to the individual bird's own song (BOS) playback and similar songs of other birds (Prather et al., 2008). Furthermore, the firing patterns of the auditory response to specific song sequences closely recapitulates the firing patterns while vocalizing the same sequences (Prather et al., 2008), suggesting that HVC_X neurons function as a primary sensorimotor integration site for vocal signals. To explore the neural representation of syntax, we focused on HVC_X neurons, and we report that these neurons code hierarchical structures of syntax with syllable-selective and transition-selective burst spikes.

Materials and Methods

All procedures were in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals* and were approved by the Institutional Animal Care and Use Committee of Kyoto University.

Subjects. Adult male Bengalese finches and zebra finches (>130 d post-hatch) were used for experiments. Bengalese finches were raised in our aviary (14:10 h light:dark cycle) and transferred to sound-attenuating chambers during the recording of their songs. Zebra finches were reared using a protocol previously described (Funabiki and Konishi, 2003; Funabiki and Funabiki, 2008). A 2.8 s Bengalese finch song sequence (a gift from Dr. Y. Funabiki, Kyoto University Faculty of Medicine, Kyoto, Japan) was used for the tutoring session for zebra finches.

Microdrive implantation surgery for electrophysiological studies. All electrophysiological studies were performed using Bengalese finches. Animals were initially anesthetized with intramuscular injection of ketamine and xylazine. A motorized microdrive and stimulus electrodes were implanted into one hemisphere (five into the left and four into the right). The microdrive was made based on the original design as previously described (Fee and Leonardo, 2001; Hahnloser et al., 2002) with some modifications. In brief, the microdrive system included three motorized positioners and a headstage preamplifier. Each positioner was equipped with custom-made micromotors (Namiki Precision Jewel) and the motion of each electrode was controlled independently. After skin incision, small skull windows were made over area X, RA, and HVC using stereotaxic coordinates. A small hole (200–300 μ m) was made in the dura over each area and bipolar stimulus electrodes were inserted to the proper depth in area X (all animals) and RA (eight of nine animals), and fixed with dental acrylic. Another small hole was made in the dura directly over HVC. The position and boundary of HVC were confirmed by the auditory response to the individual BOS. Subsequently, the exposed brain surface was protected with 2% low melting agarose in PBS and then the microdrive was implanted and secured to the skull using dental acrylic.

Electrophysiology. After complete recovery with sufficient vocalization of songs (>100 bouts/d), electrophysiological recordings were performed. During recording, microdrive electrodes were carefully advanced into HVC (5–15 μ m/step). Antidromic stimulation (Swadlow, 1998; Hahnloser et al., 2002; Fee et al., 2004) was delivered to the stimulus electrodes in area X or RA (single monopolar pulses of 200 μ s duration, 0.5 Hz, 30–150 μ A). HVC_X neurons were identified by their fixed-latency action potential responses to stimulation in area X. All the identified HVC_X neurons had no response to stimulation in RA. When a single unit of neural activity had successfully been isolated and identified, we first conducted recordings of motor-related activity ($n = 9$ birds) during unprompted vocalization. After recording in the vocal phase, we next recorded sensory-related activity in the presence of BOS playback as an auditory stimulus ($n = 4$ birds). During the recording of the auditory response, behavioral states were carefully monitored using a video camera, and the arousal level of each bird was evaluated by its opened eyes and movements. When the animal was frequently vocalizing, flying, hopping, or eating, we suspended the recording in the auditory phase. The BOS (and conspecific songs and white noise if possible) were replayed with >10 s intervals at 70 dB through a speaker placed 5 cm away from the cage in the soundproof recording chamber. Electrophysiological data were amplified and bandpass-filtered between 0.3–3 kHz. Both electrophysiological and sound data were digitized (sampling rate, 50 kHz) and stored. All spike trains were confirmed to be single unit by complete suppression of the spike autocorrelation functions within ± 1 ms. After

recording sessions, electrode placement was verified via electrolytic lesions made by passing a DC current (10 μ A for 20 s) through recording electrodes; birds were deeply anesthetized with 300 μ l of 20% urethane and perfused transcardially with saline and then 4% paraformaldehyde. The electrode placements were verified in histological brain sections.

Data analysis. Acoustic features of vocal signals were discerned using spectrograms bandpass filtered between 0.5–8 kHz. Transition diagrams of individual Bengalese finch song sequences were constructed based on >100 song renditions.

Raster plots and histograms of action potential activity were constructed by aligning action potentials to the relevant syllable onset (5 ms bin size). A burst was defined as the spike activity whose instantaneous firing rate exceeded 100 Hz (Hahnloser et al., 2002). Among the total of 120 transition segments of successive syllable pairs where burst events were observed in 47 HVC_X neurons, two segments (1.7%) contained two burst events during the silent period and the vocalization of the target syllable. In these two segments, the relative latency was measured from the onset of the first burst spikes and the firing rate was calculated from all spikes, including both of the two burst spikes.

In the statistical analysis of the neural activity in the alternative transition segments, firing rates of the individual burst spikes were defined as the average of reciprocals of interspike intervals, calculated from spikes in a time window of 100 ms. For the analysis of the firing rate, the median of the time window was determined from the time of a peak in the histogram of action potential activity. The difference in firing rates among the observed alternative trajectories within six segments upstream or downstream was analyzed using the Kruskal–Wallis ANOVA test. Neuronal reliability of burst events was defined as the probability of activity per the same time window used in the analysis for the firing rate (Prather et al., 2008). When the neuronal reliability in a nonprimary transition segment was <0.3, the transition-selective activity was classified as an all-or-none type. Otherwise, the transition-selective activity was classified as an intermediate type. The ratio of firing rate or neuronal reliability between primary and nonprimary transitions was <0.1 in all the all-or-none type transition-selective activities. In all statistical analyses, the minimum significance level was set at $p < 0.05$.

To evaluate the phasic activity in consecutive repetitions of the same syllables, differences in firing rate among the cycles were examined using the Kruskal–Wallis ANOVA test. Since the total numbers of repetitions were variable, the cycle numbers were counted from both the beginning and the end. The first and last cycles were excluded in these statistical analyses because burst spikes in the first or last cycles were unstable in some cases. Phasic activity that showed no significant differences was regarded as phasic bursts with a constant firing rate during the repetition. When phasic bursts displayed significant differences in firing rate, we further examined the relationship between firing rate and cycle numbers. As a result, all of the phasic bursts displaying significant differences in firing rate showed a significant correlation between firing rate and cycle numbers. In all statistical analyses, the minimum significance level was set at $p < 0.05$.

Results

Bengalese finches vocalize variable songs consisting of acoustic elements called syllables, which are punctuated by periods of silence and have distinct acoustic properties from each other (Woolley and Rubel, 1997). Variable song sequences of the individual birds, composed according to finite-state type syntax (Okanoya, 2004), can be described using transition diagrams (Okanoya and Yamaguchi, 1997; Sakata and Brainard, 2006; Wohlgemuth et al., 2010). As illustrated in Figure 1A, a transition diagram consists of a finite number of nodes (circles A–J) and node-connecting arrows that indicate the identities of syllables and the possible transition directions between syllables, respectively. Syntactic rules that generate diversity in sequences are described as “alternative transitions” to or from syllables: convergent transitions from multiple different syllables to a junction syllable (junction transition) and divergent transitions from

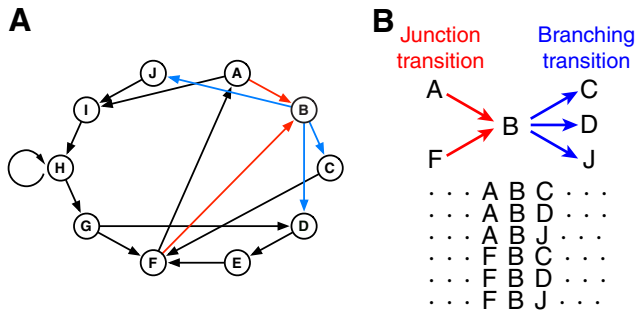


Figure 1. Alternative transitions between syllables generate song variations. **A**, An example of a syllable transition diagram representing syntactic rules. Circled letters and arrows represent the identities of syllables and the transition directions between syllables, respectively. Red and blue arrows indicate convergent transitions to syllable B and divergent transitions from syllable B, respectively. **B**, Top, Alternative transition associated with syllable B (shown in **A**) that forms a junction transition segment with syllables A and F (red) and a branching transition segment with syllables C, D, and J (blue). Bottom, Sequence variations generated from the junction and branching transitions shown above.

syllable B also forms a branching transition segment with syllables C, D, and J (Fig. 1A, blue arrows), syllable B can be followed by syllable C, D, or J (Fig. 1B, top). When a transition diagram contains no alternative transition and is composed of deterministic linear transitions, each song rendition becomes stereotypic. Thus, Bengalese finches generate complex and variable song sequences by recurrently taking different trajectories in junction and branching transition segments between syllables (Fig. 1B, bottom). To explore the neural representation of syntax for generating variable sequences, we investigated the activity profiles of single HVC_X neurons, projecting to the basal ganglia area X, which is critical for the development and maintenance of learned vocalization (Scharff and Nottebohm, 1991; Brainard and Doupe, 2000).

HVC_X activity in variable vocal sequences

We identified HVC_X neurons by antidromic activation from area X (Fig. 2A) and recorded their single-unit activity ($n = 48$ neurons, nine Bengalese finches). All HVC_X neurons except one exhibited intermittent burst spikes during singing. Although syllable sequencing of Bengalese finch song is highly variable, spike trains of HVC_X neurons displayed phase-locked patterns while vocalizing song sequences with the same syllable arrangement (Fig. 2B). To systematically evaluate the temporal relationship of the burst spikes to the associated sequences in variable songs, we divided song sequences into transition segments of successive syllable pairs and measured relative latencies (Fig. 2C). Because previous studies have demonstrated that singing-related activity in HVC_X neurons transmits motor-related information (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Prather et al., 2008) and precedes the vocal events by approximately several tens of milliseconds (Kozhevnikov and Fee, 2007), the relative latencies were defined as follows: when a burst was observed during a silent period between syllables, the relative latency was measured from the onset of the target syllable just preceded by the silent period to the onset of the burst (Fig. 2C, blue tick marks); if a burst was generated while vocalizing the target syllable, the relative latency was measured from the onset of the target syllable to the onset of the burst (Fig. 2C, red tick marks). Identified burst spikes were generated with a relative latency ranging from -100 to 100 ms. However, the relative latencies of burst events in the individual neurons grouped by their associated transition segments ($n = 120$ segments) were distributed in a more concentrated manner (average SD of the relative latency in each transition segment, 4.41 ms) (Fig. 2D), suggesting that each HVC_X neuron transmits the information related to a particular transition segment in variable vocal sequences. Twenty-three percent of HVC_X neurons ($n = 11$) exhibited burst spikes only in single transition segments (Fig. 2D, black dots, arrows), whereas the remaining 77% of the neurons

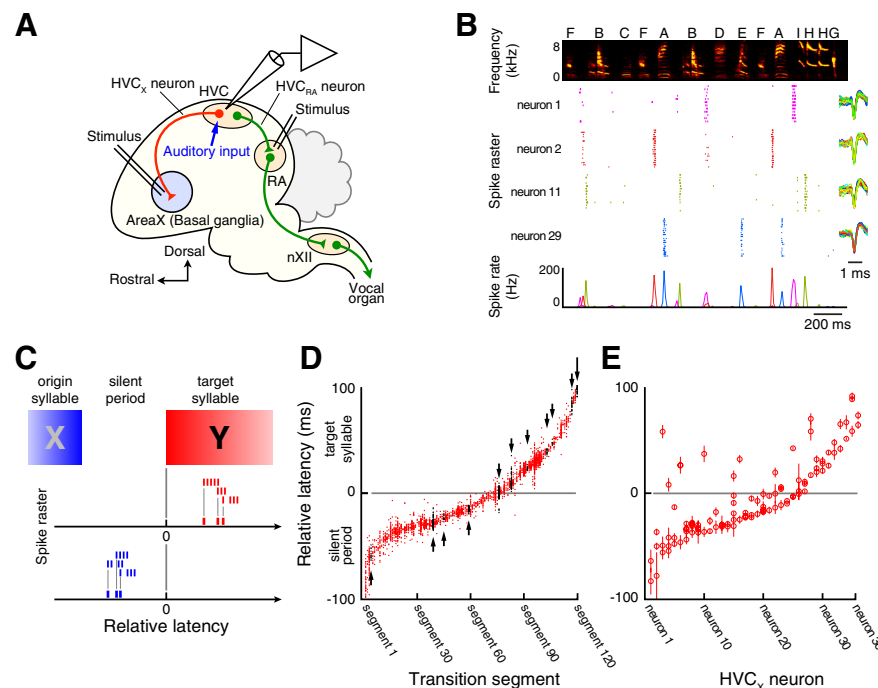


Figure 2. Activity of identified HVC_X neurons while singing syntactic vocal sequences. **A**, Schematic of a parasagittal view of Bengalese finch brain. Single-unit recordings were made in telencephalic nucleus HVC. HVC_X neurons were identified by antidromic stimulation in basal ganglia area X and RA. nXII, Nucleus XII. **B**, HVC_X neural activity recorded from a bird while vocalizing song sequences based on its own syntax (Fig. 1A). Top, Time-frequency spectrogram of the song segments. Letters above the spectrogram indicate syllable types. Middle, Spike raster plots of four HVC_X neurons. Overlaid spike waveforms of each neuron are shown at right. Bottom, Spike rate histogram illustrating burst spikes in HVC_X neurons. Tick marks representing individual spikes in the raster plots and traces in the histogram of the same neuron are colored using the same color coding. **C**, The relative latency of a burst event in the transition segment was measured from the onset of the target syllable (Y) to the onset of the burst. When the burst was generated during the origin syllable (X), the latency was measured in the preceding transition segment. **D**, The relative latencies of burst events in the individual neurons grouped by the associated segments (1–120) were sorted by their mean values and plotted. Black dots indicated by arrows identify burst events in HVC_X neurons that displayed activity only in single segments. **E**, HVC_X neurons that displayed activity in multiple segments (neurons 1–36) have preferences in spike onset latencies. The averaged relative latencies in the associated segments were grouped by the same neurons and sorted. Circles represent the averaged relative latencies to the related syllables. Error bars represent \pm SD.

a branching syllable to multiple different syllables (branching transition). For example, in Figure 1, syllable B forms a junction transition segment with syllables A and F (red arrows), and therefore can follow syllable A or F (Fig. 1B, top). In turn, because

information related to a particular transition segment in variable vocal sequences. Twenty-three percent of HVC_X neurons ($n = 11$) exhibited burst spikes only in single transition segments (Fig. 2D, black dots, arrows), whereas the remaining 77% of the neurons

($n = 36$) displayed burst spikes in multiple different transition segments ranging from 2 to 6 (Fig. 2E). Notably, although the associated syllables of the multiple burst events did not have obvious similarities in phonological structure, these neurons generated burst spikes at similar onset timing for the different transition segments (Fig. 2E). Such preferred timing in HVC_X activity suggests that the burst spikes may not directly participate in the transmission of the phonological information; rather, burst spikes of these neurons may be important for transmitting the temporal information of vocal signals. We thus focused our search on the relationship between the individual phase-locked burst spikes and syntactic information.

Syllable selectivity and transition selectivity

Assuming that the burst spikes of HVC_X neurons code syntactic structures, the individual burst events may correspond to specific syllable types or specific transition directions, because the syntactic structures can be described with a transition diagram consisting of these two types of syntactic constituents (Fig. 1A). First, we tested the possibility that HVC_X neurons can exhibit syllable-selective burst spikes in variable sequences. Figure 3A illustrates the activity profiles of neuron 11 in Figure 2B while vocalizing the junction and branching transition segment including syllable B, according to the syntactic rules represented in the transition diagram of Figure 1A. Neuron 11 preserved the identical spike pattern with the same relative latency to the onset of syllable B independently of the adjacent syllable types (Fig. 3A). Furthermore, we considered the effect of the upstream syllable arrangement in remote segments, because syllable sequencing in a branching transition segment is not a simple stochastic process in which the probability distribution of the next transition is determined by the syllable at the branching point. For example, when syllable B followed syllable A, the bird produced a primary transition from syllable B to syllable C in 52.1% of cases (Fig. 3B). In contrast, when syllable B followed syllable F, the transition probabilities were dramatically changed: the probability of the transition from syllable B to syllable C was reduced to 9.8% (Fig. 3B). Thus, HVC_X neuronal activity may be affected by the upstream syllable arrangement in remote segments. We therefore analyzed the effect of the upstream syllable sequencing on the firing rate of the burst spikes associated with syllable B. Within six transition segments upstream, three different junction transitions (Fig. 3C, black lines)

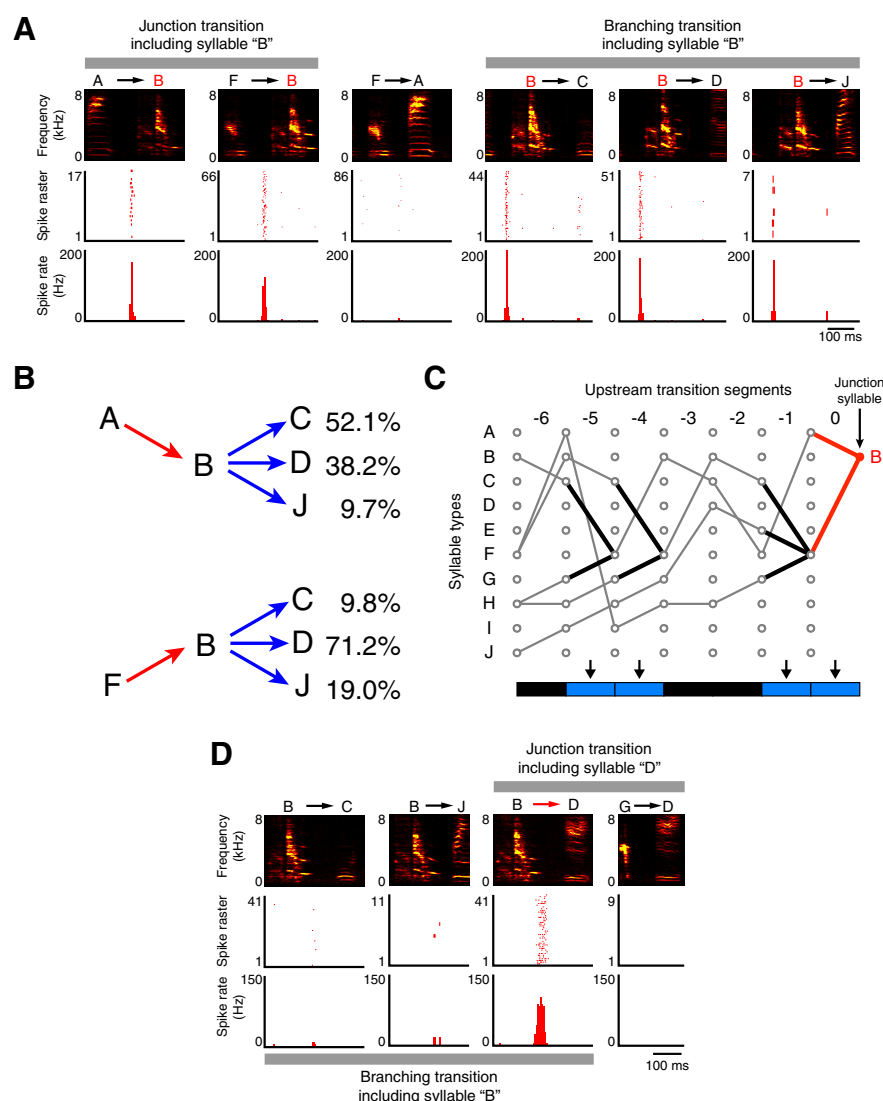


Figure 3. Syllable- and transition-selective activities, recorded from the bird that generated vocal sequences according to the syntactic rules illustrated in Figure 1A. **A**, Syllable-selective activity recorded from neuron 11 in Figure 2B. Top, Spectrograms. Middle and bottom, Raster plots and histograms, respectively, showing the spike patterns of neuron 11 while vocalizing the corresponding transition segments. Neuron 11 exhibited bursts with no statistical difference in spike rate between the junction transitions from syllable A to syllable B and from syllable F to syllable B (Kruskal–Wallis test, $p = 0.79$), and was almost inactive when vocalizing phrase FA instead of FB. The activity had no difference in spike rate among the branching transitions from syllable B to syllables C, D, and J ($p = 0.96$). **B**, The probability distribution of syllable transitions is altered depending on the sequence history. Transition probabilities of the individual trajectories are indicated on the right. **C**, Contextual differences of up to six transitions upstream had no effect on the burst spikes of neuron 11, generated in the segment including syllable B. Three different junction transitions (black lines) formed alternative trajectories toward the junction transition segment including syllable B (red line). The effects of the difference in upstream trajectories on the spike rate were statistically analyzed. The p value among alternatives was plotted using color in a rectangular cell below the corresponding junction transition segment (arrow). Blue, $p \geq 0.05$; black, linear transition segment. Note that not all of the possible syllable sequences composed according to the transition diagram (Fig. 1A) were observed, because the transition probabilities change depending on the sequence history, as shown in **B**. **D**, All-or-none type transition-selective activity recorded from neuron 1 in the same animal as in **A**. According to the transition diagram (Fig. 1A), syllable transition from syllable B to syllable D is one of the alternative transitions including syllable B (branching syllable B and the junction syllable D). Top, Spectrograms. Middle and bottom, Raster plots and histograms. Activity profiles are statistically different in rates between junction transitions BD and GD ($p = 3.4 \times 10^{-5}$) and among branching transitions BC, BD, and BJ ($p = 1.2 \times 10^{-10}$).

formed alternative trajectories toward syllable B. Statistical analysis for the effect of the difference in upstream trajectories on the firing rate revealed that the upstream syllable sequence had no effect on the activity profiles ($p \geq 0.05$) (Fig. 3C, bottom). Together, the burst spikes in the alternative transition segment can represent the identity of a specific syllable and not the history of

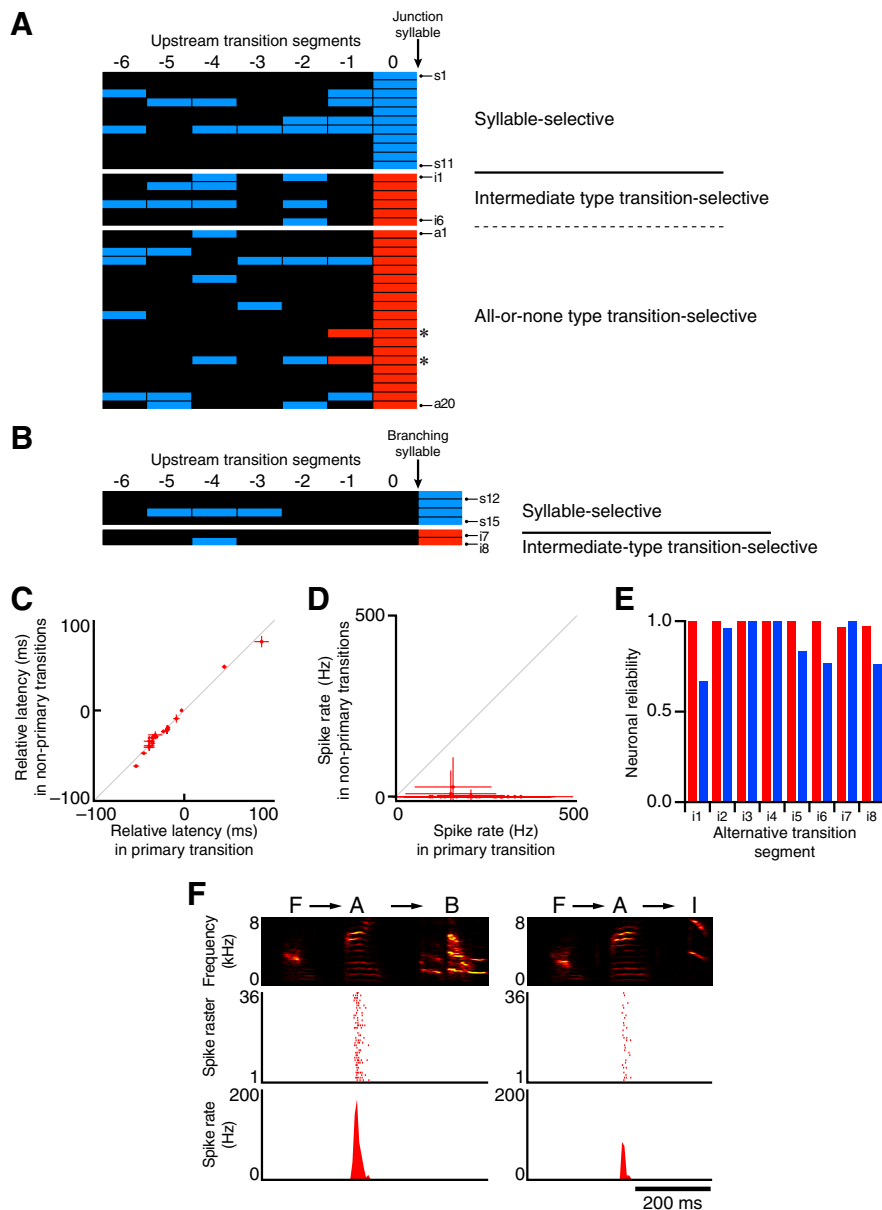


Figure 4. Singing-related HVC_x burst spikes in alternative transition segments. **A, B**, The effects of the upstream trajectories on the firing rate in the alternative transition segments were examined. Each row represents upstream transitions toward a given junction or branching syllable, where the bursts were observed. The same statistical analysis shown in Figure 3C was performed. **A**, The statistical results for the HVC_x bursts associated with junction transition segments with or without branching transitions. **B**, The results for the HVC_x bursts associated with branching transition segments without junction transitions. The lowest *p* value among alternatives was plotted using a color code (red, *p* < 0.05; blue, *p* ≥ 0.05). Black indicates linear transition segment and insufficient sampling (fewer than four samples). **C**, The syllable-selective bursts whose firing rates were constant among alternatives (s1–s15 in **A** and **B**) preserved the relative onset timings to the associated syllables independently of the adjacent syllable types. Primary or nonprimary transitions were defined by the order in the song sequences. Error bars represent ±SD. **D**, The changes in spike rate of the all-or-none type transition-selective activity between primary and nonprimary transitions (a1–a20 in **A**). Primary or nonprimary transitions were determined by the average spike rate. Error bars represent ±SD. **E**, Neuronal reliability (probability of activity per transition segment) of the intermediate type transition-selective activity (i1–i8 in **A** and **B**). Red and blue represent reliability in a primary transition and in a secondary transition, respectively. Primary or secondary transitions were determined by the average firing rate. **F**, Example of intermediate type transition-selective activity (i8 in **B** and **E**) recorded from the bird that generated song sequences according to the syntax illustrated in Figure 1A.

syllable transitions. Next, we examined the possibility of whether the burst events can represent a specific transition direction among alternatives. If HVC_x neurons display such transition selectivity, the same spike pattern should emerge only when the same transition is executed. Neuron 1, whose activity was recorded from the same animal as those in neuron 11, displayed

burst spikes only while taking the trajectory from syllable B to D among the alternatives in the branching transition segment including syllable B and in the junction transition segment including syllable D (Fig. 3D). This indicates that HVC_x neurons can also generate transition-selective burst spikes that signal a specific transition direction in alternative transition segments.

To confirm whether the burst spikes in alternative transition segments consistently represent specific syllable identities or specific transition directions, we analyzed the relationship between the syllable arrangement and the burst profiles in 43 junction or branching transition segments (*n* = 27 neurons). In 15 segments, burst spikes were generated independently of syllable patterns within six transitions upstream (Fig. 4A, s1–s15) or downstream (data not shown), similarly to those shown in Figure 3, A and C. Their average relative latencies to the associated syllables were closely matched among alternative trajectories (Fig. 4C). Thus, HVC_x neurons can consistently display syllable selectivity in variable sequences. In contrast, in 20 segments (Fig. 4A, a1–a20), firing rates changed in an all-or-none fashion (Fig. 4D), similarly to those shown in Figure 3D. In two cases where a junction or branching transition segment adjoined other alternative transition segments, the firing rate was affected by the syllable sequence in these neighboring segments (Fig. 4A, asterisks). However, alternative paths in remote segments had no effect on the firing properties, indicating that spike rate changes precisely reflect the particular trajectories in a single alternative transition segment, or two successive ones. Thus, such all-or-none type transition-selective burst spikes can explicitly signal a particular combination of two (or three) syllables in a specific order. In the remaining 18% of the alternative transition segments (i1–i8), burst spikes were observed in all transition directions and were reliable (probabilities of activity per transition segment in primary transition and secondary transitions: 0.99 ± 0.01 and 0.87 ± 0.13 , respectively; *n* = 8 cells) (Fig. 4E, F), and there existed no difference or only a subtle one in spike onset timing to the relevant syllables (Fig. 4F). Nonetheless, significant differences in spike rates among alternative

transitions distinguished this intermediate-type activity from syllable-selective activity (Fig. 4A, B). Thus, in the branching or junction transition segments, HVC_x burst spikes are categorized into three types: syllable-selective activity, all-or-none type transition-selective activity, and intermediate type transition-selective activity.

Transition dynamics in syllable repetition

Syllable repetition in the Bengalese finch is not stereotypic. In some repetition segments, the number of repeated syllables is fixed to two. However, in most of the repetition segments, the number of repetitions varies in every rendition, and such a variable repetition also contributes to producing diverse vocal patterns. Because each cycle of variable repetition can be produced according to either context-dependent or context-independent operations, we first analyzed the transition dynamics of repetitive sequences. Assuming that syllable repetition is generated simply through a repeated random process, individual transition operation between the same syllables can be regarded as an equivalent process that occurs at a constant probability. Based upon this assumption, the expected frequency distribution of syllable repeats should exponentially decrease with an increase in the number of times repeated (Fig. 5A). To test whether every transition process in a given repetition segment is actually equivalent, we examined the frequency distributions of syllable repetitions in Bengalese finches (11 segments from eight birds). Only two repetition segments displayed a monotonic decrease of distribution (Fig. 5B). All the other repetition segments (82%) displayed unimodal normal-distribution curve-like profiles (Fig. 5B). To investigate whether the transition dynamics observed in Bengalese finch songs could be generalized to bird-songs of other species, we also analyzed repetitive sequences in zebra finches (*Taeniopygia guttata*). Zebra finches generally vocalize stereotyped linear song sequences. However, they also have the ability to learn songs, including variable syllable repetition from Bengalese finch tutors (Funabiki and Konishi, 2003; Funabiki and Funabiki, 2008). We examined the frequency distributions of variable repetitions in zebra finches that successfully learned repetitive phrases consisting of Bengalese finch syllables ($n = 9$ birds) (Fig. 5D). All of the repetition segments in zebra finches displayed single-peaked, normal-distribution curve-like profiles (Fig. 5C). Thus, the assumption that repetitive transition occurs according to equivalent Markov processes does not account for the actual distribution profiles, indicating that neural states representing repetitive transitions of the same syllable change in every repetition cycle. Consequently, in cases in which spike trains with the same properties were displayed reproducibly in every repetition cycle, such activity should code the identities of repeated syllables.

Syntax-related activity in repetition

Next, we addressed whether the patterns of activity in HVC_X neurons are constant or varied during consecutive repetition in Bengalese finches. Four neurons generated burst spikes only at the first ($n = 2$ neurons) or the last ($n = 2$ neurons) cycles of the repetition phrases in an all-or-none manner (Fig. 6A,B, respec-

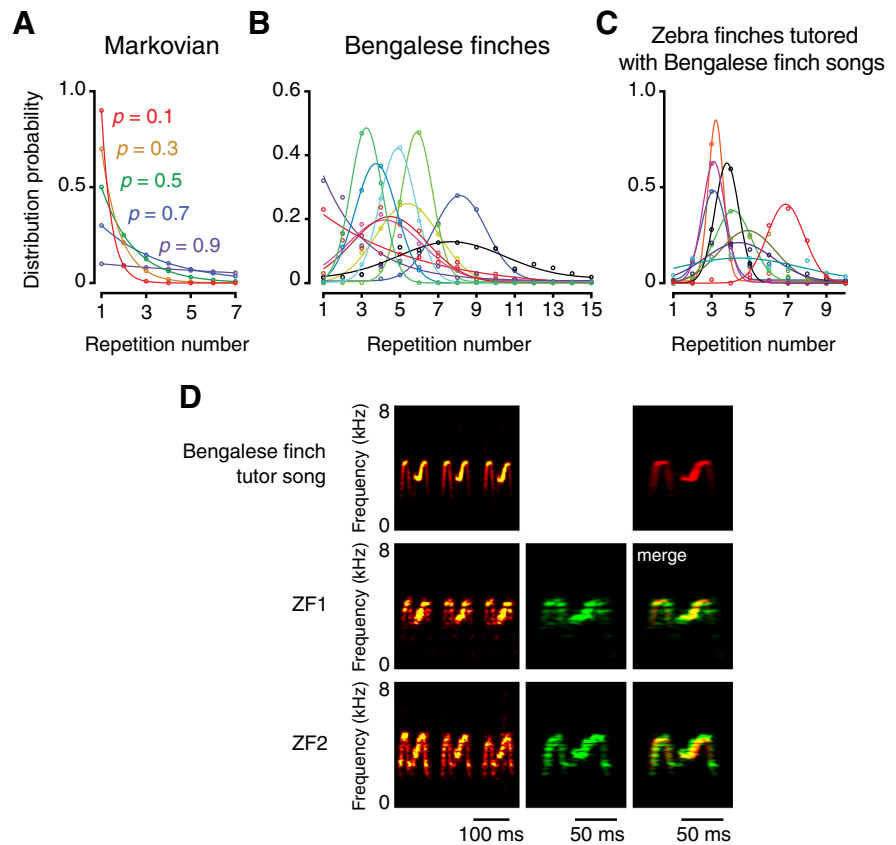


Figure 5. Non-Markovian transition dynamics in consecutive repetition. **A**, Theoretical probability distribution expected for a random process, in which transition probability between the same syllable is constant with a value of p . **B**, Actual probability distribution of syllable repeats in Bengalese finches ($n = 11$ syllables from eight birds). **C**, Probability distribution of syllable repeats in nine zebra finches that successfully learned the repetitive phrases of a Bengalese finch song. **D**, Two representative examples of consecutive repetition phrases acquired by zebra finches. Top, Spectrograms of a repetitive segment in Bengalese finch tutor song. Individual zebra finches repeated M-shaped syllables, which have never been observed in normal zebra finch song. Comparison in merged spectrograms (right) revealed acoustic and temporal similarities between the tutor syllable (red) and learned syllables (green).

tively). These all-or-none type transition-selective activities are thought to signal the initiation or termination of the syllable repetition in variable vocal sequences. We also identified a total of 20 neurons (seven birds) that generated phasic activity while vocalizing repetition segments ($n = 21$ segments). All of them exhibited brief bursts in a precisely phase-locked manner to the onset of each repetition syllable. These phasic bursts during repetition were categorized into two types: one was characterized by a constant spike rate in every repetition cycle ($n = 14$ segments) (Fig. 6C,D) and the other by a monotonically increasing or decreasing spike rate as the same syllable was repeated ($n = 7$ segments) (Fig. 6E,F). The former type preserved spike rates independently of the total number of times repeated (Fig. 6C,D, bottom right). Therefore, it represented the identities of repeated syllables as a syllable-selective activity. In contrast, the latter type represented numerical information of repetitive cycles by increasing or decreasing the spike rate during repetition and was regarded as an intermediate-type transition-selective activity. It is notable that some HVC_X neurons with an intermediate-type transition selectivity increased or decreased the firing rate adaptively to the total number of syllable repetitions, to exhibit burst spikes at a fixed firing rate every time that the first and last repetition cycle was executed (Fig. 6E,F, bottom right). Such a coding strategy is useful in indicating the progress and state-

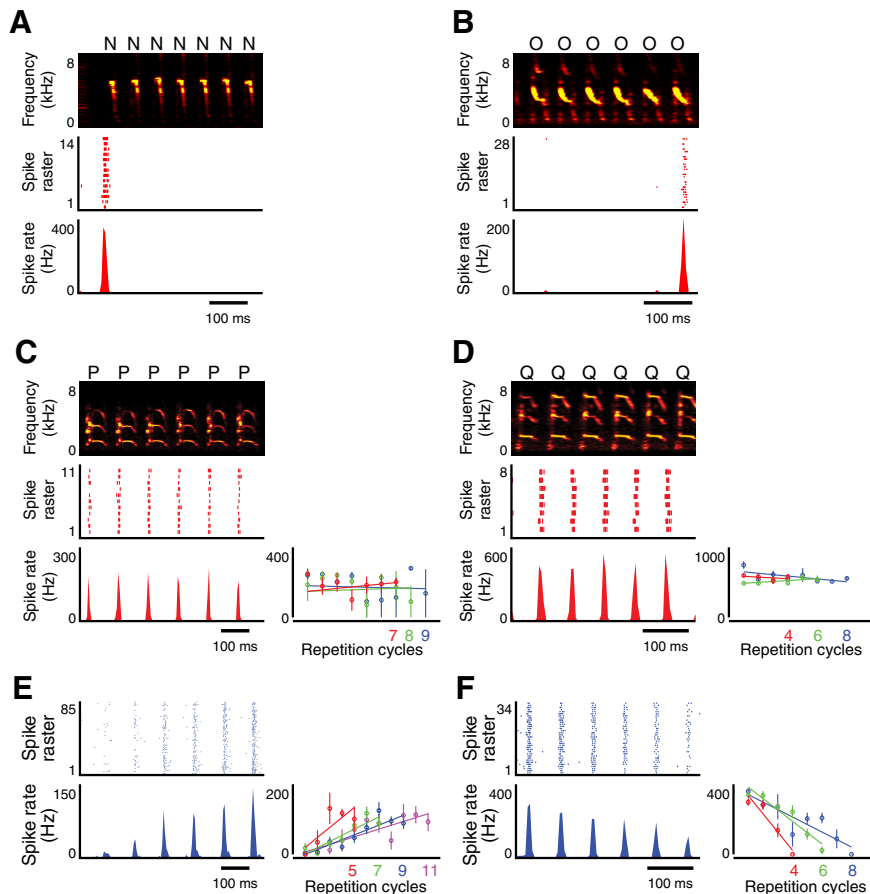


Figure 6. HVC_x neural activity while vocalizing consecutive repetitive sequences. **A, B**, All-or-none type transition-selective activity was generated only at the first cycles (**A**) or at the last cycle (**B**) while vocalizing repetitive segments. Top, middle, and bottom, Spectrograms, raster plots, and firing rate histograms, respectively. **C, D**, Top, Spectrograms in syllable repetition. Middle and bottom left, Raster plots and histograms showing syllable-selective activity, respectively. Bottom right, Firing rates of each repetition cycle along the regression lines displayed separately by the total number of repeats. Error bars represent \pm SEM. **E, F**, Top and bottom left, Raster plots and histograms showing intermediate type transition-selective activity phase-locked to the syllable repetition shown in the top rows in **C** and **D**, respectively. Bottom right, Firing rates of each repetition cycle plotted separately by the total number of repeats. Error bars represent \pm SEM.

of-completeness of the variable repetition. More importantly, these findings support the idea that the total number of repetition times is not stochastically determined during the repetition cycle, but is approximately determined in advance; that is, birds can settle on an adequate repetition number before or during the performance.

Syntactic attributes of multiple bursts in the same neurons

HVC_x bursts code hierarchical syntactic structures with syllable-selective and transition-selective activities; HVC_x neurons represent a specific syllable type and a particular combination of syllable pairs using syllable-selective activity and transition-selective activity, respectively (Fig. 7*A, B*). Furthermore, they signal the temporal boundaries of repetitive segments and the numerical information associated with repetitive cycles using all-or-none type and intermediate type transition-selective activity, respectively (Fig. 7*B*). To compute such different syntactic properties may require different types of information processing at a cellular or circuit level, leading to the idea that the difference in selectivity to specific syntactic attributes may be due to differences in cell types. To address the question of whether HVC_x neurons can be classified based on the relationship of their activity to the syntax, we examined the syntactic attributes of the

individual burst events of neurons 1–36 in Figure 2*E*, which displayed activity in multiple different segments. Single HVC_x neurons showed different syntactic properties in different segments in various combinations (Fig. 7*C*). The syllable- and transition-selective properties in repetition segments were also equally integrated with those in junction and/or branching transition segments within the burst spikes of single neurons. This indicates that syntactic selectivity of a burst event is established independently of the basic cellular profiles of the individual HVC_x neurons.

Syntax-related auditory response to self-generated vocal sequences

As previously reported (Prather et al., 2008), selective auditory response to BOS or to similar song sequences in other birds' songs observed in HVC_x neurons closely recapitulates the motor-related phasic patterns. This auditory–vocal mirroring property poses the question of whether these neurons can extract syntactic information directly from auditory inputs and represent the syntactic structures in the auditory phase in a manner correspondent with those in the vocal phase. To address this question, we compared the activity profiles of these mirroring-type HVC_x neurons in alternative transition segments between the auditory and vocal phases. Because the level of wakefulness profoundly affects the auditory response in HVC (Schmidt and Konishi, 1998), the behavioral state was monitored by a video camera in a soundproof recording chamber, and auditory stimuli were carefully

delivered while birds were active. When the BOS playback was presented as an auditory stimulus, 31% of HVC_x neurons (nine of 29 neurons, four birds) preferentially exhibited phase-locked brief activity in some of the transition segments, where the same neurons generated burst spikes during singing, and spontaneous firing of the remaining neurons in the nonsinging state was unchanged or suppressed. Although spike rates of these nine HVC_x neurons in the sensory phase were typically lower than in the motor phase, the onset timings of the auditory-related activity to the relevant segments were approximately similar to those of the vocal-related activity (Fig. 8*A*).

Among the identified nine auditory–vocal mirroring HVC_x neurons, four neurons generated bursts in the alternative transition segments in response to the BOS playback. Unlike vocal-related activity, it is logically counterintuitive that auditory transition-selective responses in branching transition segments are generated before a target syllable is presented. We therefore considered the onset timing of the bursts relative to the syllable onset. Two of these four neurons generated motor-related transition-selective bursts after the onset of the target syllables in the associated alternative transition segments (Figs. 4*A, B*, a13 and i8; 8*A*, red square and red triangle). Both of them changed their burst profiles in the segment in response to the auditory

presentation of the BOS, similar to the motor phase (Fig. 8B–D, G, H). Such sensorimotor correspondence in transition selectivity suggests that these auditory–vocal mirroring neurons extracted not only phonological but also syntactic information from auditory signals of naturally composed vocal sequences. The remaining two neurons exhibited transition-selective bursts before the onset of the target syllables in the vocal phase (Figs. 4A, a8 and a12; 8A, red inverted triangle and red circle). In the a12 segment, the burst spikes lost their transition selectivity in the auditory phase and were observed in all transition directions (Fig. 8E, I). In contrast, in the a8 segment, transition-selective changes in firing rate and neuronal reliability were preserved in the auditory phase (Fig. 8F, J). However, in the a8 segment, the HVC_X neuron exhibited bursts with slightly greater latency in the auditory phase than in the vocal phase. As a result, the auditory response was generated after the onset of the target syllable (Fig. 8A). Such a temporal shift suggests that the information processing that generates the transition-selective activity may be different depending on the respective onset timing in the transition segments. Although auditory–vocal mirroring HVC_X neurons cannot perfectly reproduce transition selectivity in the auditory phase as they do in the motor phase, our observations indicate that at least some auditory–vocal mirroring HVC_X neurons can directly extract syntactic information from sensory inputs and represent it as they do in the motor phase.

Discussion

Our results demonstrate that burst spikes in HVC_X neurons represent hierarchical syntactic structures in variable vocal sequences. HVC_X bursts during the vocalization of nonstereotyped sequences display syllable selectivity or transition selectivity (Fig. 7). In alternative transition segments between different syllables, transition-selective HVC_X bursts indicate a specific transition direction among alternatives. In consecutive repetition segments of the same syllables, all-or-none type transition-selective bursts signal entry into or exit from the repetitive sequence, and intermediate type transition-selective bursts indicate the progress and state-of-completeness of the repetition (Fig. 7B). The combination of different types of HVC_X bursts represents unique internal state changes in the course of generating variable vocal sequences. Our results also suggest that syntactic attributes of the individual HVC_X neurons are not determined by their basic cellular types but acquired during the maturation of HVC circuitry (Fig. 7C). Finally, we show that auditory–vocal mirroring type HVC_X neurons can display transition

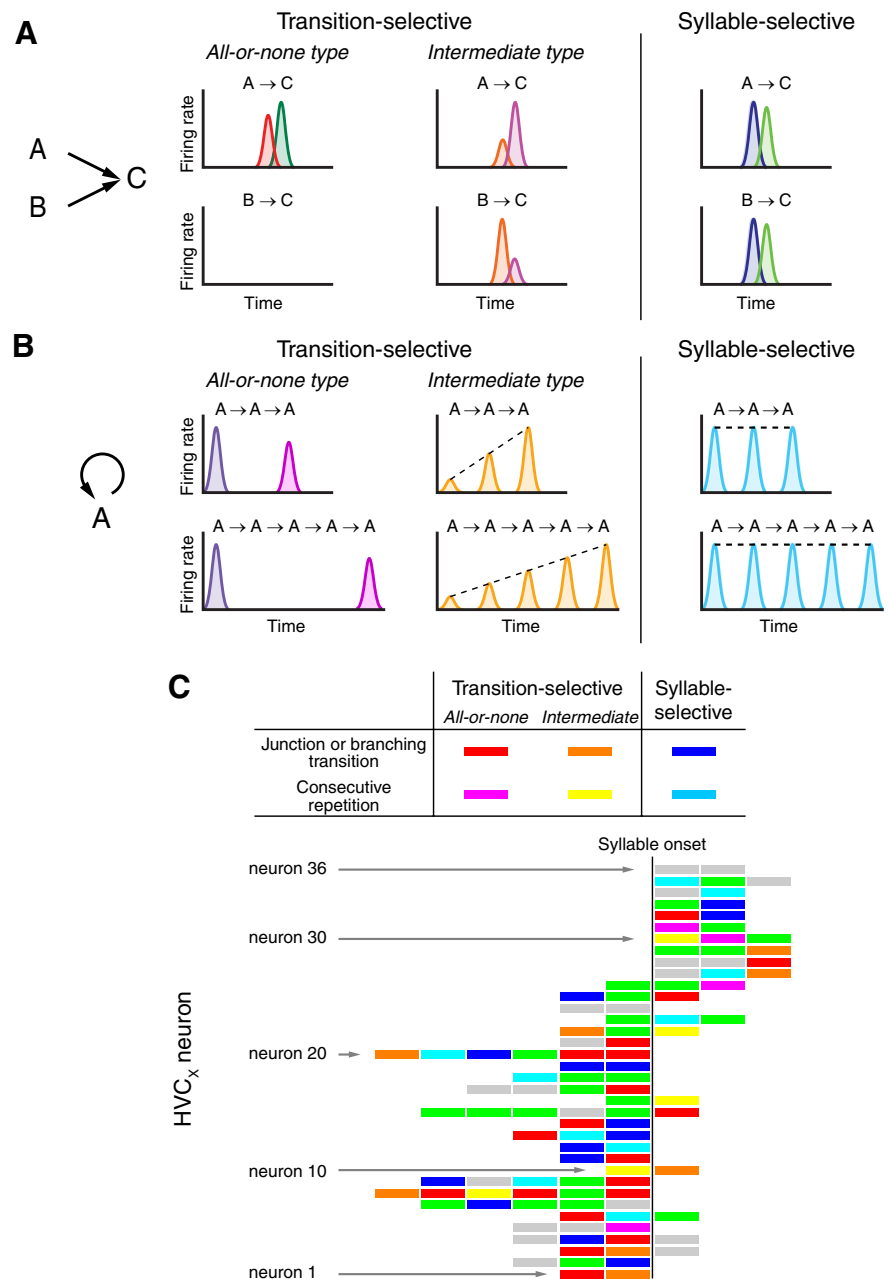


Figure 7. Neural representation of syntax in HVC_X burst spikes. **A, B**, Schematics of activity patterns of HVC_X neurons in alternative transitions between different syllable types (junction or branching transition) and in consecutive repetitions of the same syllable type, respectively. Colors for the firing rate of each HVC_X neuron are consistent in **A** and **B**. **C**, The bursts for different segments recorded in the same neuron were independently assigned to syntactic properties. Individual rectangles represent the burst events in different transition segments of neurons 1–36 shown in Figure 2E, sorted by their relative latencies. Green, Burst spikes in deterministic linear transition segments; gray, burst spikes in introductory segments or insufficient sampling.

selectivity in response to the auditory presentation of syntactic song sequences as they do in the vocal phase (Fig. 8).

Neural representation of learned behavioral sequences

In addition to vocal signaling analogous to that in avian species, mammalian subjects also learn to create complex and diverse behavioral sequences from combinations of elemental movements. In the primate supplementary motor area, some neurons exhibit activity shortly before and during the execution of a specific movement in the course of sequential tasks (Tanji and Shima, 1994) and others become active during the transition from a specific movement to another specific one (Tanji and

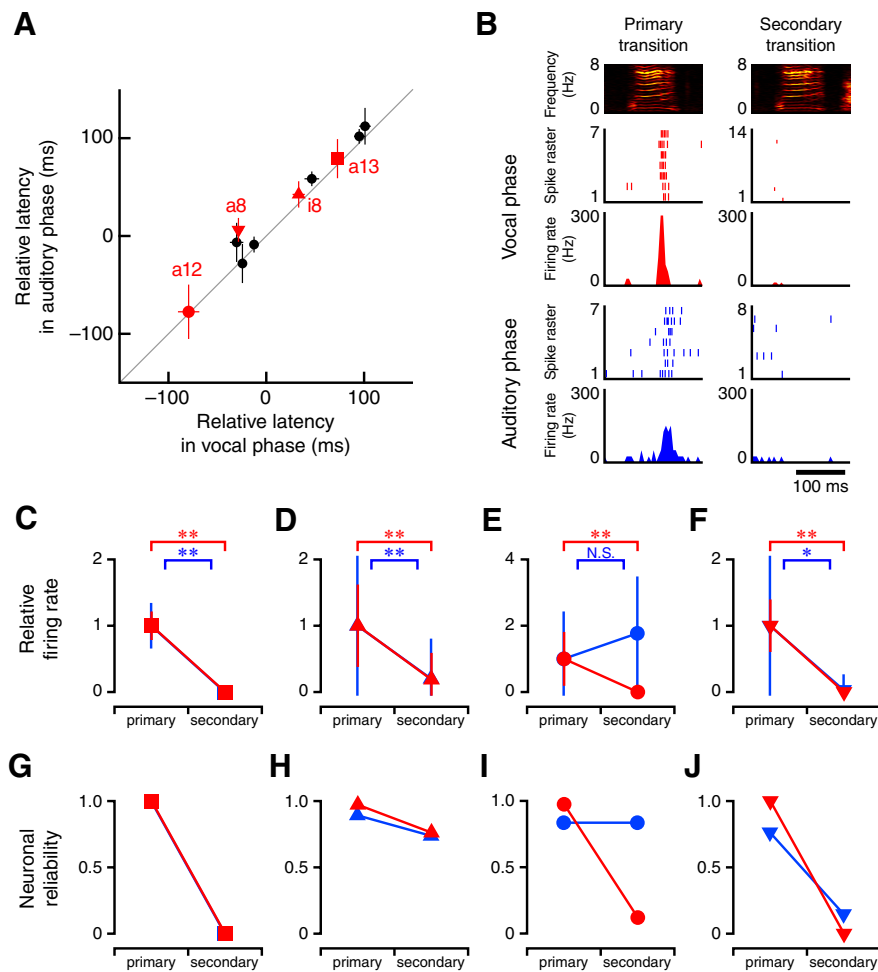


Figure 8. Transition selectivity in auditory–vocal mirroring HVC_x neurons. **A**, In 29 HVC_x neurons, nine neurons exhibited robust phase-locked activity to specific syllables (10 syllables) in response to the B05 playback (neuronal reliability in time window of 100 ms, width >0.75). The mean timing of auditory activity to the associated syllables was similar to the mean timing of motor-related activity. Error bars represent \pm SD. Red symbols indicate activities of the neurons that displayed auditory response in junction or branching transition segments where the same neurons were confirmed to generate burst spikes during singing. **B**, Example of all-or-none type transition selective activity of auditory–vocal mirroring HVC_x neuron (a13 in **A**). **C–J**, Comparison of the firing properties in the alternative transition segments between the vocal and auditory phase. Relative firing rates shown in **C–F** were normalized to the mean value in primary transitions. Primary or secondary transitions were determined by the average firing rate in the vocal phase. The same symbols as in **A** represent the same activities. Red and blue indicate activities in the motor phase and those in the auditory phase, respectively. Error bars in **C–F** indicate \pm SD. * p < 0.05; ** p < 0.001. N.S., Not significant.

Shima, 1994), similar to avian HVC_x neurons displaying syllable and transition selectivity, respectively. Thus, to represent the precise ordering of elemental actions, neural systems use transition-selective activity that codes a particular combination of movements in a specific order, in addition to movement-selective activity.

The organization of behavioral sequences also requires the proper initiation and termination of the sequence. Neurons in the primate prefrontal cortex, intimately connected with the striatum, exhibit enhanced firing at the start and end of a sequential saccade task (Fujii and Graybiel, 2003). Dopaminergic neurons in the rodent nigrostriatal pathway signal the initiation and termination of self-paced action sequences (Jin and Costa, 2010). Furthermore, our study demonstrates that basal ganglia projecting HVC_x neurons explicitly indicate the initiation and termination of syllable repetition segments, even in continuous vocal behaviors (Fig. 6A,B). Together, the hierarchical organization of learned behaviors may require the cortico–basal ganglia circuits to signal the temporal boundaries of

particular domains in the process of generating variable behavioral sequences.

To control behavioral sequences with repetitive elements, animals must settle on an adequate repetition number before or during the performance. Neurons in the primate parietal cortex display number-selective activity while performing numerically based behavioral tasks (Sawamura et al., 2002). This number-selective activity is classified as early-, late-, or middle-trial selective, according to its temporal preference. During the syllable repetition, HVC_x neurons signal the progress and state-of-completeness of the repetition with intermediate-type transition-selective bursts similar to early- or late-trial selective activity in primates (Fig. 6E,F). Such numerical representation may be useful for monitoring execution of individual actions and keeping track of the number of actions performed, to select an appropriate forthcoming action (Sawamura et al., 2002).

Because the mammalian subjects described above were all trained with intensive operant conditioning, the neural mechanisms of their learning process may differ from those of a natural learning process of imitation, such as vocal learning. Nevertheless, despite the differences in learning process, motor-related activity in avian subjects and that in mammalian subjects during the execution of complex behavioral sequences share many common features. The neural coding strategy to represent syntactic rules observed in these animal subjects may have implications for human advanced skills, including languages.

Redundant coding for syntactic information

HVC_x neurons code syntactic structures in a redundant manner with syllable-

selective and two types of transition-selective activities (Fig. 7). Because syllable identities can be included in transition-selective activity, one might therefore expect that syntactic information could be described simply by transition-selective activity. However, is such a redundant coding used for the organization of syntactic vocal signals?

All-or-none type transition-selective activity constitutes non-overlapping neural ensembles and can explicitly indicate a specific transition direction among the alternatives (Fig. 3D). However, if syntactic rules were described only with all-or-none type transition-selective activity, information about the same syllable in different syllable transitions would be represented by entirely different combinations of HVC_x activities. As a result, neurons in area X would have to extract information for the same syllable in different trajectories from completely different spatio-temporal patterns of HVC_x neuronal ensembles. To get around this problem, the intermediate type transition-selective activity, which is generated at the same timing at identical syllables, and

syllable-selective activity, which displays exactly the same spatio-temporal patterns at identical syllables, may play a part to establish and maintain precise phonological structures by presenting the identities of the same syllables in variable sequences (Fig. 7A).

Development of syntactic selectivity

Computing different syntactic properties may require different types of information processing at a cellular or circuit level, leading to the idea that differences in syntactic attributes observed in HVC_X neurons may be due to a difference in cell types. However, our results indicate that HVC_X neurons cannot be classified from their functional relationship to syntax (Fig. 7). Both syllable and transition selectivities were randomly intermingled within single HVC_X neurons (Fig. 7C). In contrast, each HVC_X neuron generated burst spikes at a similar latency to syllables with different phonological profiles (Fig. 2E). Such temporal preference of the individual neurons despite the lack of their association with specific syntactic or phonological information may be a reflection of the developmental process of the HVC circuitry.

Song develops gradually through plastic vocalizations toward a crystallized form. In this process, syllables of different types need not emerge from primitive versions of each of these types but originate in common prototypes (Tchernichovski et al., 2001). Therefore, HVC_X neurons may first tune their approximate spike timing to the common syllable prototype before each syllable becomes differentiated in acoustic and syntactic aspects. Subsequently, HVC_X activity may be shaped and establish syllable or transition selectivity in accordance with maturation in the phonological and syntactic profiles.

Intracellular electrophysiological studies have so far demonstrated that the burst spikes in HVC_X neurons arise through different subthreshold processes from those in HVC_{RA} neurons (Lewicki, 1996; Mooney, 2000; Mooney and Prather, 2005; Long et al., 2010): the temporal precision of HVC_{RA} bursts is mainly determined by selective excitatory synaptic inputs, whereas inhibitory synaptic inputs play primary roles in the establishment of HVC_X bursts. Because inhibitory synaptic inputs to HVC_X neurons presumably originate from interneurons within HVC, syntactic attributes in HVC_X activity may be formed through the development of inhibitory interneurons after the establishment of the excitatory network. In support of this idea, functional maturation of the inhibitory interneuronal network in the mammalian cortex is a protracted process, extending into late development (Dornn et al., 2010). However, the postnatal developmental process of both excitatory and inhibitory networks in HVC remains largely unknown. A full understanding of the development of syntactic processing will require further studies focused on functional maturation of the HVC circuitry.

Neural representation of syntax in auditory information

Imitative learning requires forming a memory of others' performance on the basis of sensory experience and adjusting motor commands using sensory feedback (Meltzoff, 1990). Therefore, the brain must establish a correspondence between sensory and motor representation of the performance. Such sensorimotor correspondence at a single neuron level was first identified in visual-motor mirroring neurons in primate premotor cortex (area F5) (Rizzolatti et al., 1996), regarded as a homolog of the human language area. In birds, auditory-vocal mirroring HVC_X neurons in the premotor-basal ganglia pathway (Prather et al., 2008) constitute a part of the vocal control system. Because vocal behavior requires both phonological and syntactic control, auditory-vocal correspondence should be established not only in

phonological but also syntactic aspects. Although auditory-vocal mirroring HVC_X neurons were less active in the auditory phase, our results demonstrate that these mirroring HVC_X neurons can display transition selectivity in response to the auditory presentation of syntactic song sequences, just as they do in the vocal phase (Fig. 8). Because HVC receives several different inputs (Nottebohm et al., 1982; Bottjer et al., 1989; Vates et al., 1996; Bauer et al., 2008), attenuation of the HVC_X activity in the auditory phase may reflect a difference in active input between the auditory and vocal phases. In addition, such an attenuation in the auditory phase may be partly due to selective attention, which modulates neural responses to sensory stimuli (Maunsell, 1995; Desimone, 1996; Newsome, 1996). In this regard, selective attention may suppress auditory responses to the playback of the BOS sequences as unattended stimuli.

Although the physiological role and significance of HVC_X activity during the auditory phase is still open to question, the fact that HVC_X neurons can directly extract syntactic information from the playback of naturally composed variable vocal sequences suggests that these songbirds may have the ability to learn syntactic rules from auditory experience and apply them to form their own vocal signals. Because syntax of the birdsong is coded by neurons that are analogous to mammalian cortical neurons projecting to basal ganglia circuits (Jarvis et al., 2005; Bolhuis et al., 2010) that are involved in the sensorimotor and cognitive processes of language (Ullman, 2001; Frey and Gerry, 2006), the principles obtained from the studies of the avian vocal control system may contribute to the further understanding of the neural basis of syntax in human language.

References

- Bauer EE, Coleman MJ, Roberts TF, Roy A, Prather JF, Mooney R (2008) A synaptic basis for auditory-vocal integration in the songbird. *J Neurosci* 28:1509–1522.
- Bolhuis JJ, Okanoya K, Scharff C (2010) Twitter evolution: converging mechanisms in birdsong and human speech. *Nat Rev Neurosci* 11:747–759.
- Bottjer SW, Halsema KA, Brown SA, Miesner EA (1989) Axonal connections of a forebrain nucleus involved with vocal learning in zebra finches. *J Comp Neurol* 279:312–326.
- Brainard MS, Doupe AJ (2000) Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* 404:762–766.
- Chomsky C (1969) The acquisition of syntax in children from 5 to 10. Cambridge: MIT press.
- Chomsky N (1965) Aspects of the theory of syntax. Cambridge: MIT.
- Desimone R (1996) Neural mechanisms for visual memory and their role in attention. *Proc Natl Acad Sci U S A* 93:13494–13499.
- Dornn AL, Yuan K, Barker AJ, Schreiner CE, Frome RC (2010) Developmental sensory experience balances cortical excitation and inhibition. *Nature* 465:932–936.
- Dutar P, Vu HM, Perkel DJ (1998) Multiple cell types distinguished by physiological, pharmacological, and anatomic properties in nucleus HVC of the adult zebra finch. *J Neurophysiol* 80:1828–1838.
- Fee MS, Leonardo A (2001) Miniature motorized microdrive and commutator system for chronic neural recording in small animals. *J Neurosci Methods* 112:83–94.
- Fee MS, Kozhevnikov AA, Hahnloser RH (2004) Neural mechanisms of vocal sequence generation in the songbird. *Ann N Y Acad Sci* 1016:153–170.
- Frey SH, Gerry VE (2006) Modulation of neural activity during observational learning of actions and their sequential orders. *J Neurosci* 26:13194–13201.
- Fujii N, Graybiel AM (2003) Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science* 301:1246–1249.
- Funabiki Y, Funabiki K (2008) Song retuning with tutor model by adult zebra finches. *Dev Neurobiol* 68:645–655.
- Funabiki Y, Konishi M (2003) Long memory in song learning by zebra finches. *J Neurosci* 23:6928–6935.
- Gentner TQ, Hulse SH, Bentley GE, Ball GF (2000) Individual vocal recog-

- tion and the effect of partial lesions to HVC on discrimination, learning, and categorization of conspecific song in adult songbirds. *J Neurobiol* 42:117–133.
- Hahnloser RH, Kozhevnikov AA, Fee MS (2002) An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* 419:65–70.
- Jarvis ED, Güntürkün O, Bruce L, Csillag A, Karten H, Kuenzel W, Medina L, Paxinos G, Perkel DJ, Shimizu T, Striedter G, Wild JM, Ball GF, Dugas-Ford J, Durand SE, Hough GE, Husband S, Kubikova L, Lee DW, Mello CV, et al (2005) Avian brains and a new understanding of vertebrate brain evolution. *Nat Rev Neurosci* 6:151–159.
- Jin X, Costa RM (2010) Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* 466:457–462.
- Katz LC, Gurney ME (1981) Auditory responses in the zebra finch's motor system for song. *Brain Res* 221:192–197.
- Kelley DB, Nottebohm F (1979) Projections of a telencephalic auditory nucleus-field L in the canary. *J Comp Neurol* 183:455–469.
- Konishi M (1965) The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z Tierpsychol* 22:770–783.
- Kozhevnikov AA, Fee MS (2007) Singing-related activity of identified HVC neurons in the zebra finch. *J Neurophysiol* 97:4271–4283.
- Kuhl PK (2004) Early language acquisition: cracking the speech code. *Nat Rev Neurosci* 5:831–843.
- Lewicki MS (1996) Intracellular characterization of song-specific neurons in the zebra finch auditory forebrain. *J Neurosci* 16:5855–5863.
- Long MA, Fee MS (2008) Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* 456:189–194.
- Long MA, Jin DZ, Fee MS (2010) Support for a synaptic chain model of neuronal sequence generation. *Nature* 468:394–399.
- Margoliash D (1983) Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J Neurosci* 3:1039–1057.
- Maunsell JH (1995) The brain's visual world: representation of visual targets in cerebral cortex. *Science* 270:764–769.
- McCasland JS, Konishi M (1981) Interaction between auditory and motor activities in an avian song control nucleus. *Proc Natl Acad Sci U S A* 78:7815–7819.
- Meltzoff AN (1990) Towards a developmental cognitive science: the implications of cross-modal matching and imitation for the development of representation and memory in infancy. *Ann N Y Acad Sci* 608:1–31, discussion 31–37.
- Mooney R (2000) Different subthreshold mechanisms underlie song selectivity in identified HVC neurons of the zebra finch. *J Neurosci* 20:5420–5436.
- Mooney R, Prather JF (2005) The HVC microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways. *J Neurosci* 25:1952–1964.
- Newsome WT (1996) Visual attention: spotlights, highlights and visual awareness. *Curr Biol* 6:357–360.
- Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in the canary, *Serinus canarius*. *J Comp Neurol* 165:457–486.
- Nottebohm F, Kelley DB, Paton JA (1982) Connections of vocal control nuclei in the canary telencephalon. *J Comp Neurol* 207:344–357.
- Okanoya K (2004) The Bengalese finch: a window on the behavioral neurobiology of birdsong syntax. *Ann N Y Acad Sci* 1016:724–735.
- Okanoya K, Yamaguchi A (1997) Adult Bengalese finches (*Lonchura striata* var. domestica) require real-time auditory feedback to produce normal song syntax. *J Neurobiol* 33:343–356.
- Prather JF, Peters S, Nowicki S, Mooney R (2008) Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature* 451:305–310.
- Rizzolatti G, Fadiga L, Gallese V, Fogassi L (1996) Premotor cortex and the recognition of motor actions. *Brain Res Cogn Brain Res* 3:131–141.
- Saffran JR (2002) Constraints on statistical language learning. *J Mem Lang* 47:172–196.
- Sakata JT, Brainard MS (2006) Real-time contributions of auditory feedback to avian vocal motor control. *J Neurosci* 26:9619–9628.
- Santelmann LM, Jusczyk PW (1998) Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition* 69:105–134.
- Sawamura H, Shima K, Tanji J (2002) Numerical representation for action in the parietal cortex of the monkey. *Nature* 415:918–922.
- Scharff C, Nottebohm F (1991) A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J Neurosci* 11:2896–2913.
- Schmidt MF, Konishi M (1998) Gating of auditory responses in the vocal control system of awake songbirds. *Nat Neurosci* 1:513–518.
- Swadlow HA (1998) Neocortical efferent neurons with very slowly conducting axons: strategies for reliable antidromic identification. *J Neurosci Methods* 79:131–141.
- Tanji J, Shima K (1994) Role for supplementary motor area cells in planning several movements ahead. *Nature* 371:413–416.
- Tchernichovski O, Mitra PP, Lints T, Nottebohm F (2001) Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291:2564–2569.
- Ullman MT (2001) A neurocognitive perspective on language: the declarative/procedural model. *Nat Rev Neurosci* 2:717–726.
- Vates GE, Broome BM, Mello CV, Nottebohm F (1996) Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. *J Comp Neurol* 366:613–642.
- Vu ET, Mazurek ME, Kuo YC (1994) Identification of a forebrain motor programming network for the learned song of zebra finches. *J Neurosci* 14:6924–6934.
- Wohlgemuth MJ, Sober SJ, Brainard MS (2010) Linked control of syllable sequence and phonology in birdsong. *J Neurosci* 30:12936–12949.
- Woolley SM, Rubel EW (1997) Bengalese finches *Lonchura striata* domestica depend upon auditory feedback for the maintenance of adult song. *J Neurosci* 17:6380–6390.
- Yu AC, Margoliash D (1996) Temporal hierarchical control of singing in birds. *Science* 273:1871–1875.